



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/075,323	02/14/2002	Sandra L. Godsey	020017	8686

23696 7590 03/14/2005

Qualcomm Incorporated
Patents Department
5775 Morehouse Drive
San Diego, CA 92121-1714

EXAMINER

SKED, MATTHEW J

ART UNIT PAPER NUMBER

2655

DATE MAILED: 03/14/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No. 10/075,323	Applicant(s) GODSEY ET AL.	
	Examiner Matthew J Sked	Art Unit 2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☐ Responsive to communication(s) filed on ____.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☐ Claim(s) ____ is/are pending in the application.
- 4a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) ____ is/are allowed.
- 6) ☒ Claim(s) 1-21 is/are rejected.
- 7) ☐ Claim(s) ____ is/are objected to.
- 8) ☐ Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 14 February 2002 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. ____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. ____. |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date <u>2/14/02</u> . | 6) <input type="checkbox"/> Other: ____. |

DETAILED ACTION

Specification

1. The disclosure and claims are objected to because the term "voice recognition" is misused for what nowadays is called --**speech recognition**-- in the speech signal processing art. While "voice recognition" and "speech recognition" were both once used interchangeably to refer to spoken word recognition, nowadays these two terms are distinguished. The term "**voice** recognition" now denotes identification of *who* is doing the speaking (class 704/246), while "**speech** recognition" (or "**word** recognition") denotes identification of *what* is being said (class 704/251). So, appropriate correction to the proper terms of art is required.

2. The disclosure is objected to because of the following informalities: the disclosure is missing a summary.

Appropriate correction is required.

Claim Objections

3. Claim 20 is objected to because of the following informalities: "phenome" should be replaced with --phoneme--.

Appropriate correction is required.

Claim Rejections - 35 USC § 102

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

Art Unit: 2655

A person shall be entitled to a patent unless –

(e) the invention was described in a patent granted on an application for patent by another filed in the United States before the invention thereof by the applicant for patent, or on an international application by another who has fulfilled the requirements of paragraphs (1), (2), and (4) of section 371(c) of this title before the invention thereof by the applicant for patent.

5. Claims 16-17 are rejected under 35 U.S.C. 102(e) as being anticipated by Franz et al. (U.S. Pat. 6,356,865).

As per claim 16, Franz teaches a unit for speech processing, comprising:
converting the input speech signal to a textual output (utterance hypothesis generator generates text from the speech and displays it to the screen, col. 16, lines 30-33); and

synthesizing the textual output using a plurality of speech parameters (speech synthesis unit would inherently have a plurality of speech parameters in order to modify the synthesis filter such that it will output a signal exhibiting the traits of speech, Fig. 3, element 312).

6. As per claim 17, Franz teaches a database for storing the plurality of speech parameters (does not teach that these speech parameters are dynamically generated so therefore they must be stored in a database, speech synthesis unit, Fig. 3, element 312).

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and

Art Unit: 2655

the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

8. Claims 1-15 and 18-21 rejected under 35 U.S.C. 103(a) as being unpatentable over Franz in view of Abe (U.S. Pat. 5,940,797).

As per claim 1, Franz teaches an apparatus, comprising:

a speech recognition unit adapted to receive a speech input and generate a textual output (utterance hypothesis generator generates text from the speech and displays it to the screen, Fig. 3, element 302 and col. 16, lines 30-33);

a speech synthesis unit coupled to the voice recognition unit, adapted to receive the textual output and generate a speech output (speech synthesis unit, Fig. 3, element 312); and

a database coupled to the speech synthesis unit, adapted to store speech parameters (does not teach that these speech parameters are dynamically generated so therefore they must be stored in a database, speech synthesis unit, Fig. 3, element 312).

Franz does not teach a training unit adapted to acquire speech samples and provide speech parameters to the database.

Abe teaches a text to speech synthesizer that has a training unit adapted to acquire speech samples and provide speech parameters to the database (auxiliary information extraction unit extracts the fundamental frequency, power and phoneme duration from speech and saves it in memory, col. 5, lines 25-31 and Fig. 1, elements 20 and 34).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Franz to have a training unit adapted to acquire speech samples and provide speech parameters to the database as taught by Abe because, as taught by Abe, recording speech messages for multiple tones and speeds would be burdensome hence it would save time to extract the speech features from the training speech (col. 1, line 63 – col. 2, line 9).

9. As per claim 2, Franz teaches the speech synthesis unit retrieves the speech parameters (speech synthesis unit would necessarily have to retrieve the speech parameters from memory in order to synthesize speech, speech synthesis unit, Fig. 3, element 312).

10. As per claims 3 and 18, neither Franz nor Abe specifically teach or suggest the speech parameters are diphones.

However, the Examiner takes Official Notice that diphones are notoriously well known in the art. Therefore it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Franz and Abe so the speech parameters are diphones because this would require a smaller memory and would be simpler to implement in concatenative synthesis methods.

11. As per claims 4 and 19, Franz does not teach the training unit is operative to modify speech parameters of the speech samples and to store the modified speech parameters in the database.

Abe teaches the training unit is operative to modify speech parameters of the speech samples and to store the modified speech parameters in the database (editors

form a GUI to allow the user to modify the prosodic parameters and stores them in memory, col. 7, lines 60-67).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Franz so the training unit is operative to modify speech parameters of the speech samples and to store the modified speech parameters in the database as taught by Abe because it would allow the user to have more control over the synthesis hence making the system more adaptable to the user's wishes.

12. As per claim 5, Franz teaches a linguistic parameter database for storing grammatical reference information and dictionary entries (Analyzer for Inflectional Morphology linguistically analyzes the text to determine inflection for synthesis, col. 21, lines 36-54).

13. As per claim 6, Franz teaches a translation coupled between the voice recognition unit and the speech synthesis unit, adapted to translate an input language into a second language (translates from source to target language, Fig. 3, element 308).

14. As per claim 7, Franz and Abe do not teach the training unit is further adapted to update the speech parameters in response to feedback based on the speech output.

However, Franz teaches using feedback to adapt the language model used for speech to text processing (Fig. 12, elements 1206, 1210, 1226, and 1228). This would suggest that adapting language models based upon feedback is well known in the art. Therefore, it would have obvious to one of ordinary skill at the art at the time of invention to use this adaptation scheme to adapt the speech parameters for the speech synthesizer because it would give an objective method to evaluate the quality of speech

and use this evaluation to improve the speech parameters to ensure a high quality of synthesized speech.

15. As per claims 8, 13 and 15, Franz teaches a method, apparatus and computer software program for speech processing, comprising:

receiving an input speech signal (Fig. 3, element 302);

converting the input speech signal to a textual output (utterance hypothesis generator generates text from the speech and displays it to the screen, col. 16, lines 30-33);

using a desired set of speech parameters (synthesizer would inherently have a set of speech parameters in order to modify a synthesis filter such that it will output a signal exhibiting the traits of speech, speech synthesis unit, Fig. 3, element 312) ; and

synthesizing the textual output using the desired set of speech parameters (speech synthesis unit, Fig. 3, element 312).

Franz does not teach selecting a desired set of speech parameters.

Abe teaches selecting a desired set of speech parameters (stores multiple prosodic parameters in memory and reads a selected one out for synthesis, col. 7, lines 1-12).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Franz to select the desired set of speech parameters for synthesis as taught by Abe because it would enable the system to synthesize text using multiple voices hence making the system more enjoyable for the user.

16. As per claims 9 and 14, Franz does not teach receiving speech samples to build a speech parameter database; extracting speech parameters from the speech samples; modifying the speech parameters to form modified speech parameters; and storing the modified speech parameters; and using the modified speech parameters to synthesize speech.

Abe teaches:

receiving speech samples to build a speech parameter database and extracting speech parameters from the speech samples (auxiliary information extraction unit extracts the fundamental frequency, power and phoneme duration from speech and saves it in memory for retrieval by the speech synthesizer, col. 5, lines 25-31 and Fig. 1, elements 20 and 34);

modifying the speech parameters to form modified speech parameters (editors form a GUI to allow the user to modify the prosodic parameters and stores them in memory, col. 7, lines 60-67); and

storing the modified speech parameters; and using the modified speech parameters to synthesize speech (parameters provided to the speech synthesizer, col. 8, lines 18-21).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Franz to have a training unit for extracting speech parameters from speech signals, modifying the speech signals and storing and using the speech parameters to synthesize speech as taught by Abe because it would enable

the system to synthesize text using multiple voices hence making the system more enjoyable for the user.

17. As per claim 10, neither Franz nor Abe teach that modifying the speech parameters comprises comparing the speech samples to a target speech sample and removing the irregularities from the speech samples.

However, the Examiner takes Official Notice that modifying parameters to match a good voice is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Franz and Abe to compare the speech samples to a target speech sample and remove the irregularities from the speech samples because this would allow the current speech to be synthesized to be modeled after a desired good voice hence ensuring good synthesized speech quality.

18. As per claim 11, Franz does not teach that extracting speech parameters comprises identifying speech units within the speech samples.

Abe teaches identifying speech units within the speech samples (determines the start and end points of phonemes, Fig. 2, element 25A).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Franz to identify speech units within the speech samples as taught by Abe because partitioning speech into phonemes is well known method to synthesize speech from its most basic components.

19. As per claim 12, Franz and Abe do not teach receiving feedback information based on application of the speech output, determining an accuracy of the application of

Art Unit: 2655

the speech output and if the accuracy is less than a predetermined threshold, updating the modified speech parameters.

However, Franz teaches using feedback to adapt the language model used for speech to text processing (Fig. 12, elements 1206, 1210, 1226, and 1228). This would suggest that adapting language models based upon feedback is well known in the art. Therefore, it would have obvious to one of ordinary skill at the art at the time of invention to use this adaptation scheme to adapt the speech parameters for the speech synthesizer because it would give an objective method to evaluate the quality of speech and use this evaluation to improve the speech parameters to ensure a high quality of synthesized speech.

20. As per claim 20, Franz does not teach the speech-to-text unit provides phoneme boundary information to the training unit.

Abe teaches the training unit extracts phoneme boundary information from the training speech (determines the start and end points of phonemes, Fig. 2, element 25A).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of to use the speech-to-text unit of Franz to determine the phoneme boundary information instead of the training unit as taught by Abe because it would give a better indication of the phoneme boundary information for the current speech being analyzed, hence giving better synthesis results.

21. As per claim 21, neither Franz nor Abe specifically teach the training unit is activated during a training mode, and deactivated during a normal operating mode.

However, the Examiner takes Official Notice that training is performed prior to regular operation of a system is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Franz and Abe because it would prohibit the training mode from interfering with the operation of the system and would save memory requirements, hence speeding up the system.

Conclusion

22. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Bakis et al. (U.S. Pat. 6,859,778), Greene, Jr. et al. (U.S. Pat. 6,377,925), and Brittan et al. (U.S. Pat. Pub. 2002/0184030A1) teach alternative methods for speech-to-text/text-to-speech translators. Chu et al. (U.S. Pat. Pub. 2002/0099547A1), Tischer (U.S. Pat. Pub. 2004/0111271A1), Savic (U.S. Pat. 5,327,521), Gibson et al. (U.S. Pat. 6,336,092) and Case et al. (U.S. Pat. Pub. 2002/0193995A1) teach methods for modifying the synthesized voice based upon trained speech segments.

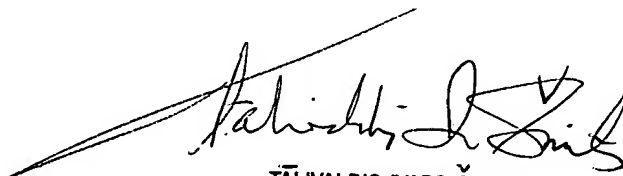
Any inquiry concerning this communication or earlier communications from the examiner should be directed to Matthew J Sked whose telephone number is (703) 305-8663. The examiner can normally be reached on Mon-Fri (8:00 am - 4:30 pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on (703) 306-3011. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Art Unit: 2655

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

MS
03/01/05



TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER